# Mathematical Statistics

Kenichi ISHIKAWA

http://ishiken.free.fr/lecture.html

Oct. 18 Combination and probability
Oct. 25 Random variables and probability distributions
Nov. 1 Representative probability distributions
Nov. 8 (First half) Random walk and gambler's ruin problem
(Latter half) Brownian motion and diffusion
Nov. 22 Noise theory

Only exercises are provided on Nov. 15.

# Reference books


[1] Satsuma, J. (2001). *"Probability /Statistics"* -- *Beginning course of mathematics of science and technology 7*, Iwanami Shoten

[2] Kolmogorov, A.N., Žurbenko, I. G., Prokhorov, A.V. (2003). *"Beginning of Kolmogorov's Probability Theory"*. (Trans. Murayama , T. & Baba, Y.): Morikita Shuppan

[3] Kitahara, K. (1997). *"Nonequilibrium Statistical Mechanics"* – *Iwanami Fundamental Physics Series 8*, Iwanami Shoten

# Mathematical Statistics

Kenichi ISHIKAWA

http://ishiken.free.fr/lecture.html

Nov. 11 Representative probability distributions

- Binomial distribution
- Poisson distribution
- Normal distribution
- Central limit theorem

# 3-1 Binomial Distribution

→ Reference book [1], p.68

- Definition of binomial distribution

[Example] Cast a die five times, supposing that the number of times of getting one-spot on the die is the random variable *X*. Then, what probability distribution will *X* follow?

Number when $X=2$ ⇨ ${}_5C_2 = 10$ cases

Probability of occurrence of each individual case ⇨ $\frac{1}{6}\times\frac{1}{6}\times\frac{5}{6}\times\frac{5}{6}\times\frac{5}{6}=\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^3$

One-spot: twice

Numbers other than one: three times

Probability Density $f(2)=10\times\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^3=\frac{625}{3888}=0.161$ ← Probability of getting one-spot twice

$$f(0)=\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^5=\frac{3125}{7776}=0.402 \qquad f(1)=5\times\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^4=\frac{3125}{7776}=0.402$$

(5 = ${}_5C_1$)

## • Definition of binomial distribution

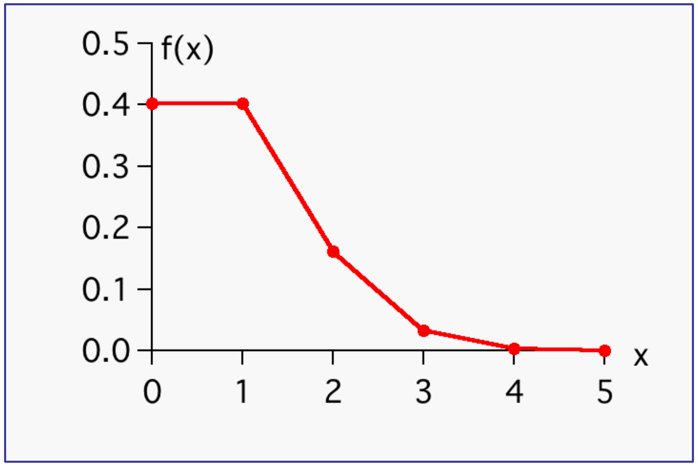
[Example] Cast a die five times, supposing that the number of times of getting one-spot on the die is the random variable X. Then, what probability distribution will X follow?

$$f(2) = 10 \times \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^3 = \frac{625}{3888} = 0.161$$

$$f(0) = \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^5 = \frac{3125}{7776} = 0.402$$

$$f(1) = 5 \times \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^4 = \frac{3125}{7776} = 0.402$$

$$f(3) = \frac{125}{3888} \qquad f(4) = \frac{25}{7776} \qquad f(5) = \frac{1}{7776}$$




**Binomial Distribution（Bernoulli distribution）**

Given that the probability of occurrence of an event $P(A) = p$, the probability of $x$-times occurrence of $A$ throughout $n$-times independent trials is:

$$Bin(n, p) \Rightarrow f(x) = {}_nC_x\, p^x (1-p)^{n-x} \qquad (x = 0, 1, 2, \ldots, n)$$

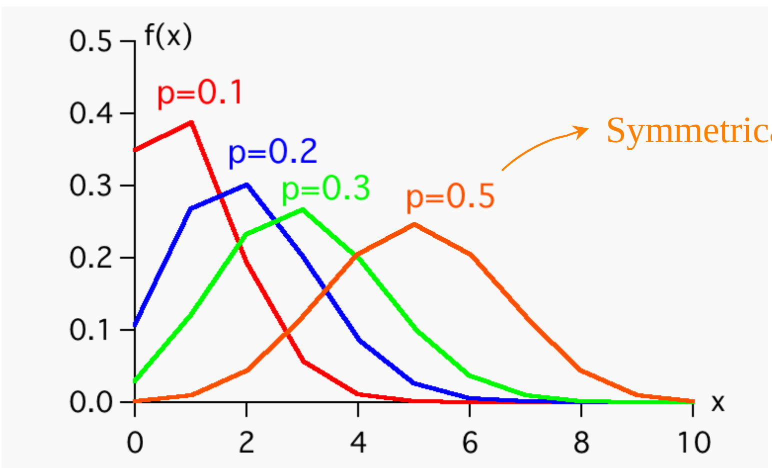
- [Examples] Suppose that there are 10 five-choice questions, each of which is allocated 10 marks. What is the probability of getting 80 marks or more when answering these questions based entirely on guesswork ?

→ Reference book [1], p.69

$Bin(10,1/5)$ ⇨

$$f(8)+f(9)+f(10)={}_{10}C_8\left(\frac{1}{5}\right)^8\left(\frac{4}{5}\right)^2+{}_{10}C_9\left(\frac{1}{5}\right)^9\left(\frac{4}{5}\right)^1+{}_{10}C_{10}\left(\frac{1}{5}\right)^{10}\left(\frac{4}{5}\right)^0$$

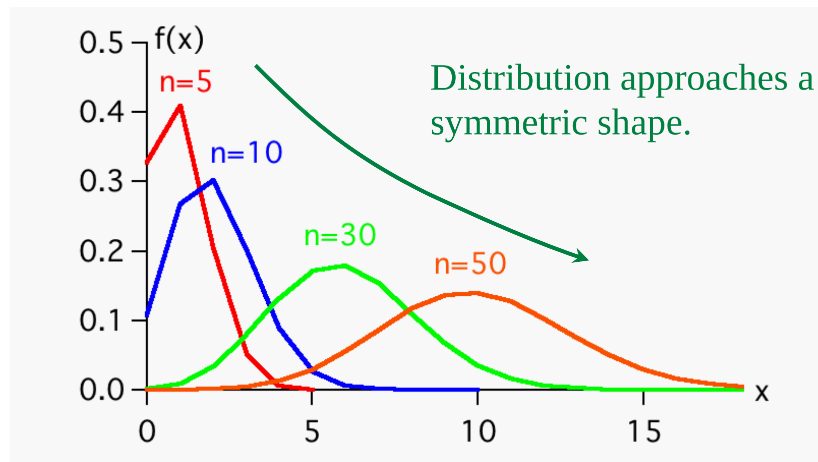
$$=\frac{761}{9765625}=0.000078$$




Binomial Distribution

$Bin(10, p)$

Correct answer rate for each question




Binomial Distribution

$Bin(n,0.2)$

## 3-1 Binomial Distribution

- Properties of binomial distribution → Related to binomial theorem

$$f(x) = {}_nC_x p^x (1-p)^{n-x} \xrightarrow{q=1-p} f(x) = {}_nC_x p^x q^{n-x}$$

→ Each term of a binomial expansion formula of $(p+q)^n = \sum_{x=0}^{n} {}_nC_x p^x q^{n-x}$

$$(p+q)^n = \sum_{x=0}^{n} {}_nC_x p^x q^{n-x}$$

→ $\sum_{x=0}^{n} f(x) = 1$

Differentiated by $p$ →

$$n(p+q)^{n-1} = \sum_{x=0}^{n} x\,{}_nC_x p^{x-1} q^{n-x}$$

Multiplied by $p$ →

**Expectation value**

$$\mu_x = np$$

Differentiated by $p$ →

$$n(n-1)(p+q)^{n-2} = \sum_{x=0}^{n} x(x-1)\,{}_nC_x p^{x-2} q^{n-x}$$

Multiplied by $p^2$ →

$$n(n-1)p^2 = \sum_{x=0}^{n} (x^2 - x) f(x)$$

→

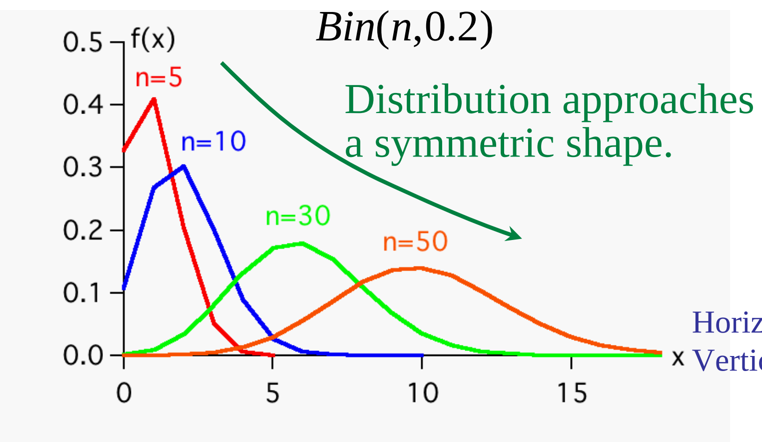
$$n(n-1)p^2 + \mu_x = \sum_{x=0}^{n} x^2 f(x)$$

→

**Variance**

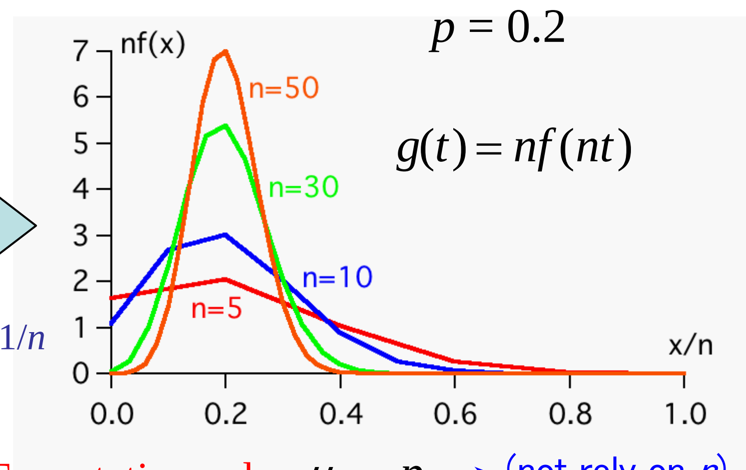
$$\sigma_x^2 = n(n-1)p^2 + np - n^2p^2$$
$$= np(1-p)$$

- Law of Large Numbers




$$T = \frac{X}{n}$$

Horizontal axis times $1/n$
Vertical axis times $n$




Expectation value $\mu_x = np$

Variance $\sigma_x^2 = np(1-p)$

Expectation value $\mu_t = p$ → (not rely on $n$)

Variance $\sigma_t^2 = \frac{p(1-p)}{n} \xrightarrow[n\to\infty]{} 0$

→ Reference book [1], p.74

→ Important base for mathematically handling empirical probability !

- Law of large numbers
  - You can only probabilistically tell whether an event $A$ occurs in each single trial. However, the greater the number of times of trials, the closer the ratio of occurrence of the event $A$ comes to a certain value $p$.

# 3-2 Poisson distribution

→ Reference book [1], p.75

Binomial Distribution $f(x) = {}_nC_x p^x (1-p)^{n-x}$

The limit of $n \to \infty, p \to 0$ is taken, keeping the mean $\mu = np$ at a fixed value.

$$f(x) = \frac{n(n-1)\cdots(n-(x-1))}{x!}\left(\frac{\mu}{n}\right)^x \left(1-\frac{\mu}{n}\right)^{n-x}$$

$$= \frac{n^x}{x!}\cdot 1 \cdot \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\left(\frac{\mu}{n}\right)^x \left(1-\frac{\mu}{n}\right)^{n-x}$$

$$= \frac{\mu^x}{x!}\cdot 1 \cdot \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\left(1-\frac{\mu}{n}\right)^n \left(1-\frac{\mu}{n}\right)^{-x}$$

$$= \frac{\mu^x}{x!}\cdot 1 \cdot \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\left[\left(1-\frac{\mu}{n}\right)^{-n/\mu}\right]^{-\mu} \left(1-\frac{\mu}{n}\right)^{-x}$$

$$\xrightarrow[n\to\infty]{} \frac{\mu^x}{x!}e^{-\mu}$$

Poisson distribution $f(x) = \dfrac{\mu^x}{x!}e^{-\mu}$

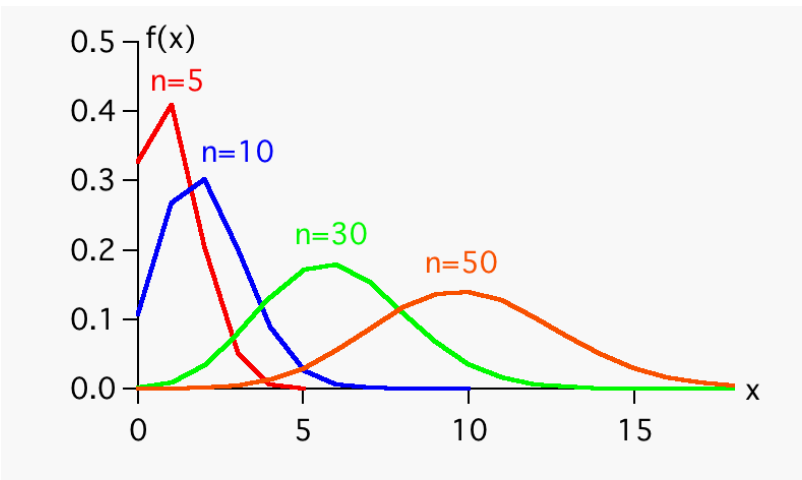
## 3-2 Poisson distribution

Poisson distribution $P(\mu)$ $f(x)=\frac{\mu^x}{x!}e^{-\mu}$

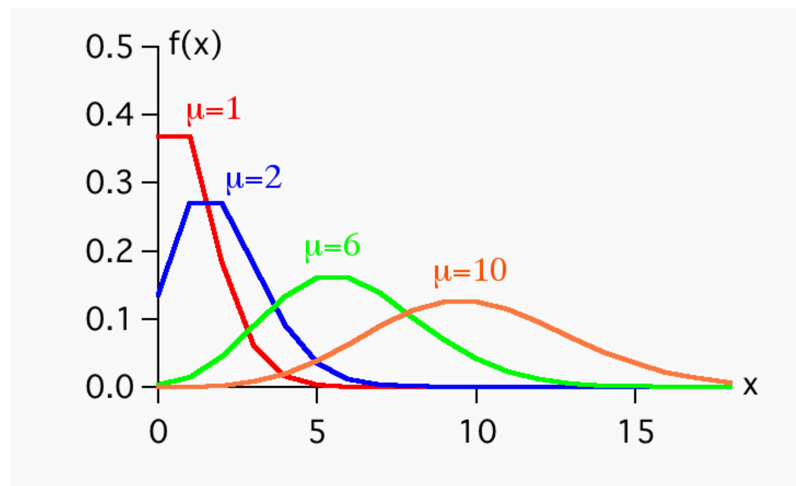
- Events having a lower probability of occurrence ($p$ is small.)
- Multiple independent trials ($n$ is large.)

Variance $\sigma^2 = np(1-p)=\mu\left(1-\frac{\mu}{n}\right)\xrightarrow[n\to\infty]{}\mu$

$\sigma^2=\mu,\ \sigma=\sqrt{\mu}$




Binomial distribution $Bin(n,0.2)$




Poisson distribution $P(\mu)$

東京大学工学部 システム創成学科 Cコース Program for Biological & Information Systems

## 3-2 Poisson distribution

Poisson distribution $P(\mu)$ $f(x)=\frac{\mu^x}{x!}e^{-\mu}$

[Examples] Number of Prussian soldiers kicked to death by horses over a period of 20 years from 1875 to 1894 in Prussia（Source: Reference book [1], p.76）

Rarity

| Number of deaths | 0 | 1 | 2 | 3 | 4 | Total |
|---|---|---|---|---|---|---|
| Number of troop units | 109 | 65 | 22 | 3 | 1 | 200 |

$$\mu=(0\times 109+1\times 65+2\times 22+3\times 3+4\times 1)/200=0.61$$

Theoretical values in the case of the Poisson distribution $P(0.61)$ are tabulated as follows:

| Number of deaths | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| Number of troop units | 108.7 | 66.3 | 20.2 | 4.1 | 0.6 |

## 3-2 Poisson distribution

Poisson distribution $P(\mu)$ $$f(x)=\frac{\mu^x}{x!}e^{-\mu}$$

[Examples] How many coming-from-behind occurred per game among 1,560 regular season games of both the Central and Pacific League of the Nippon Professional Baseball (NPB) within a given two years?（Source: Reference book [1], p.77）

Rarity

| Number of occurrence of coming-from-behind | 0 | 1 | 2 | 3 | 4 or more | Total |
|---|---|---|---|---|---|---|
| Frequency | 944 | 457 | 128 | 25 | 6 | 1,560 |

Average $\mu=0.52$

Theoretical values in the case of the Poisson distribution $P(0.52)$ are tabulated as follows:

| Number of occurrence of coming-from-behind | 0 | 1 | 2 | 3 | 4 or more |
|---|---|---|---|---|---|
| Frequency | 927.0 | 482.5 | 125.6 | 21.8 | 3.1 |

## 3-2 Poisson distribution

Poisson distribution $P(\mu)$ $f(x) = \frac{\mu^x}{x!} e^{-\mu}$

- Events having a lower probability of occurrence ($p$ is small.)
- Multiple independent trials ($n$ is large.)

Rare events occurring among extremely large number of people and matters

- Radioactive decay constant per minute of a radioisotope (Radiation count)
- Number of traffic accidents per day
- Number of plane crash accidents per year
- Number of palpable earthquakes per month

[Examples] Suppose that four customers visit a “ramen” shop every 10 minutes on average. Find the probability that six or more customers visit this ramen shop in 10 minutes.

Suppose that there are no customers visiting the shop in group.

Here, let us consider that the number of customers $X$ *follows the* Poisson distribution $P(4)$.

$$1 - \sum_{x=0}^{5} f(x) = 1 - \left( \frac{4^0}{0!} + \frac{4^1}{1!} + \frac{4^2}{2!} + \frac{4^3}{3!} + \frac{4^4}{4!} + \frac{4^5}{5!} \right) e^{-4} = 1 - \frac{643}{15} e^{-4} = 0.21$$

## 3-2 Poisson distribution

Time interval distribution of an event following the Poisson distribution

[Examples] Where a radioisotope decays once per minute on average (Radiation is counted once per minute), what will the time interval distribution g(t) of two successive decays (count) be ?

The probability of no count for *t* minutes after a count $p(t)$ is:

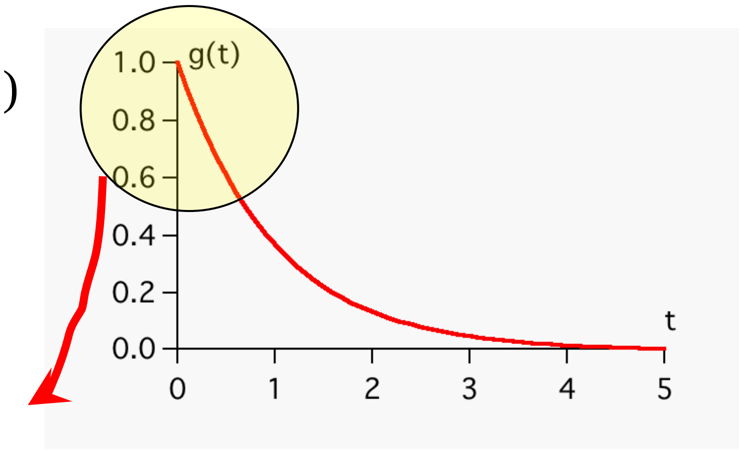
$$p(t) = e^{-t}$$

Meanwhile;

$$\int_t^{\infty} g(t')dt' = p(t)$$

→

$$g(t) = -p'(t) = e^{-t}$$

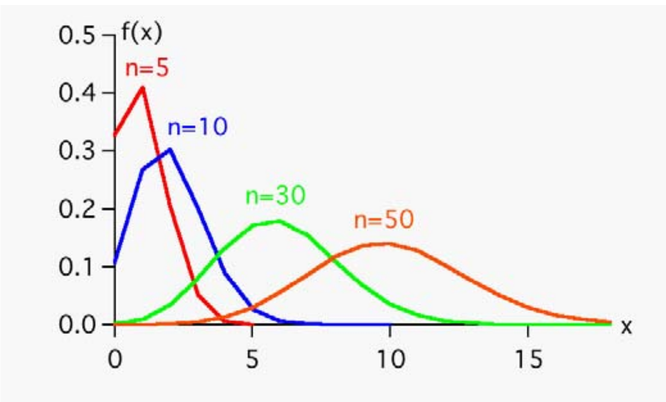



Events that occur independently of each other are -- contrary to intuition -- likely to occur in a row.

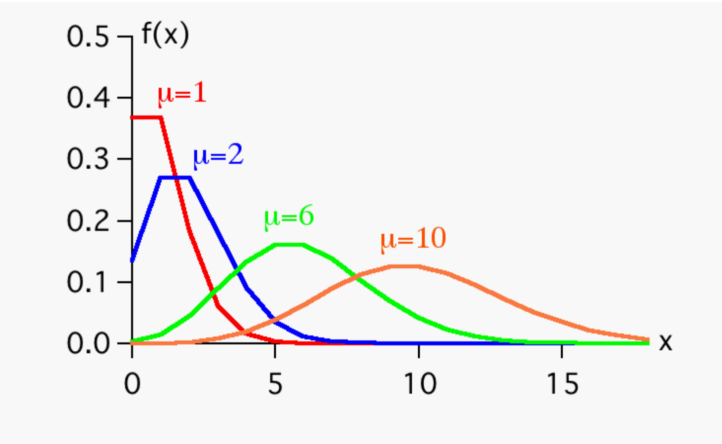
# 3-3 Normal Distribution

→ Reference book [1] , p.82

- What distribution pattern will the binomial distribution show in the large $n$ limit?
- What distribution pattern will the Poisson distribution show in the large $\mu$ limit?




Binomial Distribution $Bin(n,0.2)$




Poisson distribution $P(\mu)$

Normal Distribution

By equating $Z$ and $\frac{X-\mu}{\sigma}$ , $Z$ follows

$$g(z)=\frac{1}{\sqrt{2\pi}}e^{-z^2/2}$$

Mean 0, Variance 1

Standard normal distribution

$N(0,1)$

# 3-3 Normal Distribution

Standard normal distribution $N(0,1)$

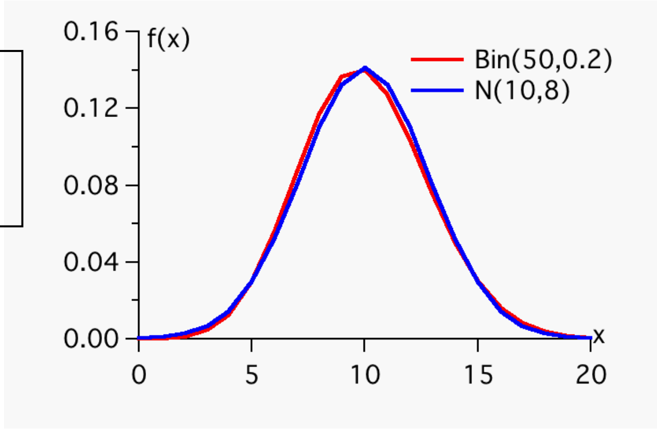
$$g(z)=\frac{1}{\sqrt{2\pi}}e^{-z^2/2}$$

Variable transformation (Standardized transformation)

Normal Distribution $N(\mu,\sigma^2)$

$$h(y)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(y-\mu)^2}{2\sigma^2}\right]$$

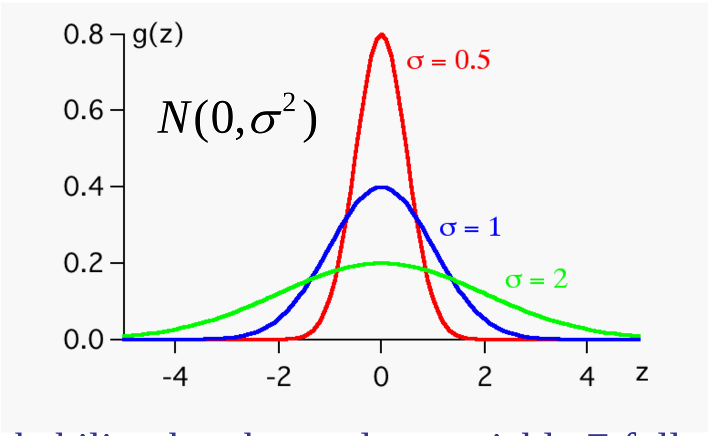
A.k.a. the Gaussian distribution




Transition from a binomial distribution to a normal distribution → An example of the Central Limit Theorem

- Various distributions approximately follow the normal distribution such as:
  - Height, weight, and exam marks
  - A wide range of science and engineering domains including experimental errors
- Binomial distribution → Starting point of various probability distributions
  Normal distribution → Most important in practice

## Properties of normal distribution




- Symmetrical
- The greater the standard deviation σ is, the flatter the distribution curve is.

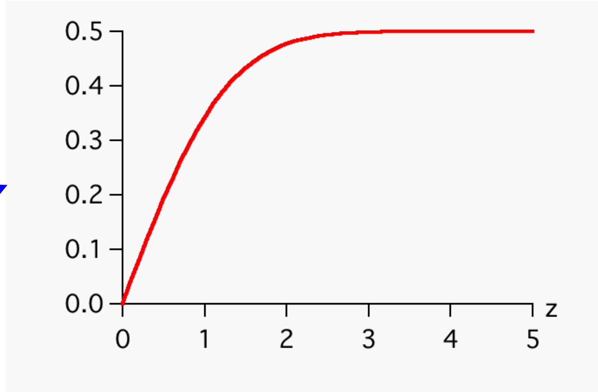
Probability that the random variable Z following the standard normal distribution $N$ (0,1) stays within the range of $z_1 < Z < z_2$

$$P(z_1 < Z < z_2) = \int_{z_1}^{z_2} g(z)dz = \int_{z_1}^{z_2} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

Error Function erf(z)

$$\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt = 2\int_0^{\sqrt{2}z} g(t)dt$$

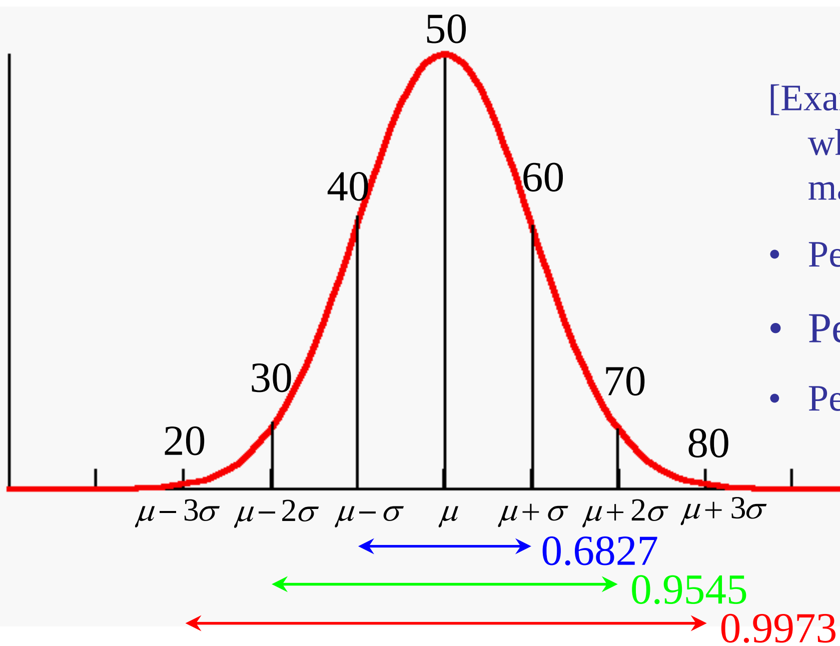
$$\Rightarrow \int_0^z g(t)dt = \frac{\mathrm{erf}(z/\sqrt{2})}{2}$$



## Probability frequently used in practice for normal distribution $N\ (m,\ s^2)$

Random variable $Y$ follows the normal distribution $N(\mu, \sigma^2)$ ⇨ $Z = \dfrac{Y-\mu}{\sigma}$ follows the normal distribution $N(1, 0)$.

$$P(\mu-\sigma < Y < \mu+\sigma) = P(-1 < Z < 1) = \mathrm{erf}(1/\sqrt{2}) = 0.6827$$
$$P(\mu-2\sigma < Y < \mu+2\sigma) = P(-2 < Z < 2) = \mathrm{erf}(2/\sqrt{2}) = 0.9545$$
$$P(\mu-3\sigma < Y < \mu+3\sigma) = P(-3 < Z < 3) = \mathrm{erf}(3/\sqrt{2}) = 0.9973$$




[Examples] On an examination, the average mark was 60, while the standard deviation was 10. Assuming that the marks of this exam follow the normal distribution;

- People attaining 80 marks or more: 2.3%
- People attaining less than 50 marks: 16%
- People belonging to a range from 40 to 80 marks: 95%

# 3-4 Central Limit Theorem

→ Reference book [1], p. 86

Suppose that the random variables, $X_1, X_2, \ldots, X_n$, are independent of each other, and follow a critical distribution having the mean $m$ and variance$s^2$ .

For $\overline{X} = \frac{1}{n}(X_1 + X_2 + \mathcal{L} + X_n)$ , the simple average of $X_1, X_2, \ldots, X_n$,

let $Z_n = \frac{\sqrt{n}}{\sigma}(\overline{X} - \mu)$ .

Then, when increasing $n$, the distribution of $Z_n$ approaches the standardized normal distribution $N$ (0,1).

→ Reference book [1], p. 86

$$\overline{X} \approx N\left(\mu, \sigma / \sqrt{n}\right)$$

## 3-4 Central Limit Theorem

[Examples] Tossing a coin 1,000 times. What is the probability that heads comes up not less than 485 times and not more than 515 times?

The number of times heads comes up $X$ follows the binomial distribution *Bin* (1000, 1/2).

$$f(x) = {}_{1000}C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{1000-x} = \frac{1000!}{x!(1000-x)!}\left(\frac{1}{2}\right)^{1000}$$

$$P(485 \le X \le 515) = \sum_{x=485}^{515} f(x) = 0.673$$

Approximate the number of times heads comes up $X$ by normal distribution.

$$\mu = 1000 \times \frac{1}{2} = 500,\ \sigma = \sqrt{1000 \times \frac{1}{2} \times \frac{1}{2}} = \sqrt{250} = 15.8$$

$$P(485 \le X \le 515) \approx P(\mu - \sigma < X < \mu + \sigma) = 0.683$$